Congratulations! You passed!
TO PASS 80% or higher

Keep Learning

GRADE
100%

Module 4 Quiz

LATEST SUBMISSION GRADE
100%

1. **Question: Which of the following statements are true?** — 1 / 1 point

- (•) Sometimes problems that seem straightforward to humans have to be broken down into smaller tasks for a machine learning system.
- ( ) Machine learning systems only ever make mistakes similar to those that humans make.
- ( ) Machine learning systems only make silly mistakes when it comes to image captioning, they're always perfect at other tasks.
- ( ) Images of bananas are hard to classify.

✓ **Correct**
Correct! We saw an example of this with the image captioning system that captioned images by first detecting words, then generating sentences, and then ranking those sentences on their likelihood.

2. **Which of the following statements are true?** — 1 / 1 point

- ( ) Machine learning systems possess common sense just like we humans do.
- ( ) Machine learning systems understand context, especially when it comes to image recognition systems.
- ( ) Machine learning systems love heavy metal music from Montreal.
- (•) Machine learning systems make mistakes that are strange, unexpected, and often different from the mistakes humans would make.

✓ **Correct**
Correct! We've seen many examples of machine learning systems making the kinds of errors that humans would likely never make.

3. **Which of the following statements are true?** — 1 / 1 point

- ( ) **Once you've trained a model, it can be used indefinitely.**
- ( ) **The distribution of your live data is always consistent.**
- ( ) **If you properly train your model, you can entirely eliminate the issue of overfitting.**
- (•) **If you train your model on one data pipeline, that does not guarantee the model will work for all data pipelines.**

✓ **Correct**
Correct! It can be useful to try to reuse as much code as possible in both the training data pipeline and the operational (live) data pipeline, but you will need to test to make sure your live data pipeline integrates well with your model.

4. **Which of the following are potential issues that can arise with live data? (Select all that apply.)** — 1 / 1 point

- [x] **There might be different distributions in live data than there was in your training data.**

Correct

Correct! Your training data is just a sample of the data out there so it is possible the live data has a different distribution (which can change over time).

- [x] The owner of a dataset might unexpectedly make a change to the data.

Correct

Correct! People have a tendency to change things without thinking about the larger impact of their changes.

- [x] The sheer volume of live data might be much larger than first expected.

Correct

Correct! Live data increases over time, and then keeps increasing. Live data may come in at a much higher speed compared to what you traditionally have with training data.

- [ ] There are no issues with live data if you've properly trained your model.

5. Which of the following statements are true? 1 / 1 point

- ( ) Standardization and normalization can never help with skewed distributions in your data.
- (•) Your data might be skewed along both the label distribution and along the feature space distribution.
- ( ) There is only one reason your data might be skewed: bad data collection.
- ( ) Algorithms make no assumptions about the distribution of the labels.

Correct

Correct! You need to understand the multiple ways in which your data can be skewed. This can be along both the y-axis and the x-axis.

6. Which of the following statements are true? 1 / 1 point

- ( ) The distribution of your data does not really matter when you have a small dataset.
- ( ) The same data collected from two different sources is guaranteed to have the same distribution.
- ( ) You should always make an informed decision on which learning algorithm to use based on your knowledge of the distribution and never rely on empirically testing different algorithms despite the distribution.
- (•) Usually when we talk about skewed distributions in machine learning, we are concerned with the feature space.

Correct

Correct! That is because a distributional shift in the feature space from your training data to your testing data puts you into a different learning regime.